

Uncertainty in detecting trend: a new criterion and its applications to global SST

Tao Lian¹

Received: 1 October 2015 / Accepted: 30 November 2016 / Published online: 28 December 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract In most parts of the global ocean, the magnitude of the long-term linear trend in sea surface temperature (SST) is much smaller than the amplitude of multi-scale internal variation. One can thus use a specific period in a much longer record to arbitrarily determine the sign of long-term trend, which is statistically significant, in regional SST. This could lead to a controversial conclusion on how global SST responded to the anthropogenic forcing in the recent history. In this study, the uncertainty in the linear trend due to multi-scale internal variation is theoretically investigated. It is found that the “estimated” trend will not change its sign only when its magnitude is greater than a theoretical threshold that scales the influence from the multi-scale internal variation. Otherwise, the sign of the “estimated” trend may depend on the period used. The new criterion is found to be superior over the existing methods when the de-trended time series is dominated by the oscillatory term. Applying this new criterion to a global SST reconstruction from 1881 to 2013 reveals that the influences from multi-scale internal variation on the sign of “estimated” linear trend cannot be excluded in most parts of the Pacific, the southern Indian Ocean and the northern Atlantic; therefore, the warming or/and cooling trends found in these regions cannot be interpreted as the consequences of anthropogenic forcing. It’s also suggested that the recent hiatus can be explained by combined uncertainty from internal variations at the interannual and decadal time scales.

Keywords Ordinary least-square · Linear trend · Internal variation · Global warming · Hiatus

1 Introduction

The long-term trend in global sea surface temperature (SST) plays a fundamental role in the climate system. Modeling studies indicated that the long-term SST change could affect tropical cyclone frequency and global precipitation. For example, Xie et al. (2010) reported that the enhanced warming over the equatorial Pacific and the Inter-tropical Convergence Zone (ITCZ) anchors increasing local rainfall, and the intensified cooling over the Southern Hemisphere could suppresses tropical cyclone development via reducing the tropical cyclone potential intensity. Cai et al. (2013) suggested that with the enhanced warming in the eastern equatorial Pacific under the future anthropogenic greenhouse gas releasing scenario, the occurrence of extreme El Niño-Southern Oscillation (ENSO) would be doubled. Identifying the underlying long-term global SST change in the recent history can undoubtedly advance our understanding of how the climate system responds to the anthropogenic forcing; it can also provide a benchmark for elucidation of model behaviors under the historical greenhouse gas forcing scenario (e.g., CMIP5 project; Taylor et al. 2012). However, since the long-term trend in global SST could be modulated by internal oscillations from decadal to multi-decadal time scales and these internal variations simulated in models are not necessarily in phase with those in the observations, one cannot directly compare the trends in models and observations over a fixed period and conclude how the model responds to the anthropogenic forcing. It is therefore necessary to explore the uncertainty

✉ Tao Lian
liantao@sio.org.cn

¹ State Key Lab of Satellite Ocean Environment Dynamics,
Second Institute of Oceanography, Hangzhou, China

due to internal variation on trend and to clarify the SST change due to the anthropogenic forcing.

By definition, the sign of the SST change due to the anthropogenic forcing should not change with the prolonging or/and shortening of the record. However, as the magnitude of the linear trend in global SST is generally much smaller than the amplitude of multi-scale internal variation, internal variation can sometimes modulate the sign of the linear trend in regional SST, masking the SST response to the anthropogenic forcing. For example, the linear trend in the zonal SST gradient over the equatorial Pacific was reported to have enhanced over the twentieth century (Luo et al. 2012). However, when using a shorter record starts from the 1950s, the zonal SST gradient over the equatorial Pacific was found to have weakened instead (Tokinaga et al. 2012). L'Heureux et al. (2013) further pointed out that the sign of the linear trend in SST in the central and eastern equatorial Pacific contains strong decadal variability. Using a pair of 100-year records since 1881, the signs of linear trend in the central southern Indian Ocean, the region north of Australia, the eastern off-equatorial Pacific, and the region off California coast can be opposite for two different periods (Fig. 8). Because of this kind of uncertainty in trend estimate, whether the resulting warming trend or/and cooling trend in these regions can be interpreted as the underlying long-term change related to the anthropogenic forcing is far from certain, even when the estimation is based on a 100-year record and when the estimated trend is statistically significant.

It is therefore necessary to find a criterion for determining the uncertainty in the sign of the “estimated” trend. If the de-trended time series is assumed to be independent or uncorrelated, the sign of the “estimated” linear trend estimated by the Ordinary Least-Square (OLS) method can be statistically verified if it passes, for example, the 95% confidence level. This straightforward criterion has been widely used in climate research (e.g., Deser et al. 2010; Luo et al. 2012; Fig. 8). However, a climate variable generally has a serial correlation (von Storch and Zwiers 1999). Assuming the de-trended time series as the first order autoregressive (AR1) process, Weatherhead et al. (1998; W98 hereafter) gave a measurement of the uncertainty of “estimated” linear trend, which depends on the autocorrelation at lag 1 and the variance of the de-trended data. Although this criterion can be easily applied (e.g., Loeb et al. 2007; Wielicki et al. 2013), any departure from the AR1 process makes the criterion in W98 perform poorly (Phojanamongkolkij et al. 2014). Leory et al. (2008; L08 hereafter), on the other hand, gave a criterion when the autocorrelations of the de-trended record at all possible time scales are taken into account. This criterion does not assume any form of the data and is theoretically versatile. However, the difficulty in estimating autocorrelation coefficients at the infinite lags

makes it difficult to be used in practice (Phojanamongkolkij et al. 2014).

The aim of this study is to provide a new criterion to scale the uncertainty in the sign of the “estimated” trend, especially for the time series dominated by the oscillatory term. In particular, we will discuss the role of multi-scale internal variation and the role of noise process in the uncertainty of the “estimated” trend, respectively. The rest of the paper is arranged as follows. In Sect. 2, we investigate the theoretical expression of “estimated” trend by the OLS method. A new concept named as the theoretical threshold is introduced to scale the uncertainty in the sign of the “estimated” linear trend related to the oscillatory term. The advantage of this criterion is given by an idealized experiment. In Sect. 3, we apply the new criterion to investigate whether the underlying long-term trend in the global SST since 1881 can be correctly estimated by the OLS method, followed by conclusions and discussion in Sect. 4.

2 Trend evaluation: effects of internal variation on linear trend estimation

2.1 W98 and L08 criteria

Let $y(t)$ be a continuous time series of a climate variable with zero mean. In W98, the linear trend model assumes $t) = Bt + N(t)$, where B is the “true” linear trend or the long-term linear trend, and $N(t)$ is the term for noise. Considering the associated scale of the fluctuating climate variable, the noise term is assumed to be the AR1 process in W98, that is, $N(t) = \rho_1 N(t-1) + \epsilon(t)$, where ρ_1 and $\epsilon(t)$ denote the autocorrelation of the noise term at lag 1 and white noise, respectively. Let $r(L)$ denote the “estimated” trend. When $\|\rho_1\|$ is not close to 1.0, the standard deviation of $r(L)$ used to measure the uncertainty of $r(L)$ can be approximated as:

$$U_{W98} \approx \sqrt{\frac{12}{L^3} \sigma^2 dt \left(\frac{1 + \rho_1}{1 - \rho_1} \right)} \quad (1)$$

where σ and dt are the standard deviation of $N(t)$ and the interval between samples, respectively, and L is the length of the time series. In the case that $\|\rho_1\|$ is close to 1.0, the standard deviation of $r(L)$ can be estimated by a more complex formula (see the appendix in W98).

The uncertainty of “estimated” trend proposed in W98 requires only two estimated model parameters (σ and ρ_1), making it easy to use. However, the noise term for the climate variable does not always follow the AR1 process. When this happens, U_{W98} cannot be used to estimate the uncertainty of “estimated” trend precisely (Phojanamongkolkij et al. 2014).

The criterion shown in L08 does not assume any form for the noise term. Instead, it takes the autocorrelation of $N(t)$ at all lags (from $-\infty$ to $+\infty$). The standard deviation of the “estimated” trend $r(L)$ given in L08 is:

$$U_{L08} \approx \sqrt{\frac{12}{L^3} \sigma^2 dt \left(\sum_{\mu=-\infty}^{+\infty} \rho_{\mu} \right)} \quad (2)$$

where ρ_{μ} and dt denote the autocorrelation of $N(t)$ at lag μ and the interval between samples, respectively, and σ is the deviation of $N(t)$.

Compared to the criterion proposed in W98, the criterion postulated in L08 is closer to the reality and theoretically versatile. However, the difficulty of estimating the autocorrelation of $N(t)$ at the infinite lag makes it difficult to be applied to finite time series (Phojanamongkolkij et al. 2014). It is noted that when the noise is white (i.e., $\rho_0 = 1$; $\rho_i = 0$ when $i \neq 0$), U_{W98} and U_{L08} are identical. In this case, as generally used in climate research, the difference between the “estimated” and the “true” linear trends will be less than $1.96U_{W98}$ at the 95% confidence level.

2.2 A new criterion

Here, we propose a new criterion for the uncertainty of “estimated” linear trend by decomposing the noise term $N(t)$ used in W98 into two parts: the part denoting the multi-scale oscillation that can be interpreted as multi-scale internal variation, and the part denoting the possible discontinuity arising from sampling errors and other unexpected fluctuations, which are independent from the multi-scale internal variation:

$$y(t) = Bt + \sum A_i \cdot \sin(\omega_i t + \phi_i) + N_t, \quad (3)$$

where B is the “true” linear trend or the long-term linear trend. A_i , ω_i and ϕ_i are the amplitude, frequency and phase of the i th oscillatory term, respectively. N_t represents the unexpected noise.

In W98, the noise term is regarded as the red noise and further assumed as the AR1 process. In our model, the part in the de-trended time series, which shows strong autocorrelation, has been explicitly expressed as the oscillatory term, so the residual part in the de-trended time series is assumed as uncorrelated in a first-order approximation. In other words, N_t in Eq. (3) is assumed as white noise. Because the “estimated” linear trend is a linear function of the terms used (Eq. 13 in the appendix), the “estimated” linear trend for Eq. (3) can be written as:

$$r(L) = r(B) + r(O) + r(N_t), \quad (4)$$

where $r(B)$, $r(O)$ and $r(N_t)$, and denote the part of the linear trend related to the “true” linear trend term, the oscillatory

term and the white noise term, respectively. From Eq. (13), we have $r(B) = B$. The analytical solution for $r(O)$ can also be derived (Eq. 14 in the appendix).

2.2.1 Without the white noise term

Let us first consider a case without the white noise term, that is, $N_t = 0$. In this case, $r(N_t)$ in Eq. (4) is also equal to zero. The “estimated” linear trend $r(L)$ can be approximated as (Eq. 15 in the appendix):

$$r(L) \cong B - \frac{6}{L^2} \sum \frac{A_i}{\omega_i} [\cos(\omega_i L + \phi_i) + \cos(\phi_i)], \quad (5)$$

where L is the record length.

From Eq. (5), it is clear that the “estimated” linear trend consists of two terms: One represents the “true” linear trend, and the other denotes the uncertainty of “estimated” linear trend due to the oscillatory term. When A_i and ω_i are given, the second term will only depend on two parameters. The first parameter is the record length. With the increase of the record length, the uncertainty decreases with the square of the record length. In this sense, the “true” linear trend can be precisely estimated if the record length is long enough. The second parameter is the phase of the oscillatory term. With a different phase of the oscillatory term at the beginning, the “estimated” trend changes nonlinearly. Nevertheless, since the magnitude of the sum of the terms in the bracket cannot be greater than 2, we have:

$$r(L) - B \leq \frac{12}{L^2} \sum \frac{A_i}{\omega_i} \quad (6)$$

From Eq. (6), it is clear that for a given amplitude, a low-frequency variation (small ω_i) will have a larger impact on the “estimated” trend than a high-frequency one (large ω_i). Equation (6) also indicates that the difference between the “true” and the “estimated” linear trends is smaller than the theoretical threshold of:

$$B_t = \frac{12}{L^2} \sum \frac{A_i}{\omega_i}. \quad (7)$$

Therefore, the “estimated” trend will not give a wrong estimation of the sign of the “true” linear trend once its magnitude exceeds the theoretical threshold. Otherwise, the sign of the “estimated” trend may depend on the phase of the oscillatory term (ϕ_i in Eq. 5).

We can also determine the record length required to guarantee the sign of the “true” trend by the OLS method. From Eq. (6), we have:

$$B - \frac{12}{L^2} \sum \frac{A_i}{\omega_i} \leq r(L) \leq B + \frac{12}{L^2} \sum \frac{A_i}{\omega_i}. \quad (8)$$

Therefore, if the “true” linear trend is given, its sign will not be incorrectly estimated by $r(L)$ when the record length L is greater than the critical length of

$$L_c = \sqrt{\frac{12}{\|B\|} \sum \frac{A_i}{\omega_i}}. \quad (9)$$

2.2.2 With the white noise term

Theoretically, any continuous time series can be decomposed by a set of oscillatory terms (for example, by Fourier basis functions), and we can use the theoretical threshold to scale the uncertainty of “estimated” trend. However, in practice the observed record is discrete and finite. One thus needs to consider the role of the white noise term N_t in Eq. (3), which denotes the part of the record that cannot be fit by the oscillatory term, in the uncertainty of “estimated” trend. From Eq. (1), when ρ_1 is set as zero, the standard deviation of $r(N_t)$ can be approximated as:

$$\delta \cong \sqrt{\frac{12}{L^3} \text{dt} \sigma^2}, \quad (10)$$

where σ denotes the variance associated with N_t . δ is referred as the “noise effect” hereafter.

As commonly used in scientific investigation, there is 95% probability that $\|r(L) - B - B_t\| \leq 1.96\delta$. Taking this and referring to Eq. (4), the uncertainty in the sign of the “estimated” trend $r(L)$ can be measured by $B_t + 1.96\delta$ when a noise term is considered. However, it should be noted that in some extreme cases (which count for 5% cases), $\|r(L) - B\|$ could be larger than $B_t + 1.96\delta$. Therefore, the uncertainty in the sign of $r(L)$ measured by $B_t + \delta$ is from the statistical point of view. Of course, if the noise effect δ is much smaller than the theoretical threshold B_t , the uncertainty in the sign of $r(L)$ due to the white noise can be neglected. In this case, the uncertainty in the sign of $r(L)$ can be deterministically scaled by B_t alone.

2.3 An idealized example

To illustrate the concept of the theoretical threshold, a simple idealized example is given here. The time series is set as:

$$y(t) = Bt + 100.0 \sin(1.0t + \varphi). \quad (11)$$

The record length is set to 50.0. Note that these parameters are selected for neat display in a figure, though using other parameters will give the same conclusion.

Figure 1 gives the “estimated” linear trend $r(L)$ as a function of the “true” linear trend. With the increase of the “true” linear trend, the “estimated” linear trend linearly increases as suggested by Eq. (5). However, their difference is always smaller than the required theoretical threshold of

0.48. It is clear that the sign of the “estimated” linear trend does not depend on the phase used as long as the magnitude of the “true” trend exceeds the theoretical threshold (green triangles). Otherwise, the sign of the “estimated” linear trend may depend on the phase of the oscillatory term. For example, taking the “true” linear trend as 0.2, one may get a positive “estimated” trend if the oscillatory phase is set to $\Phi = \frac{3\pi}{4}$, or one may get a negative “estimated” trend if $\Phi = -\frac{\pi}{4}$. Therefore, the theoretical threshold does provide a boundary for the magnitude of the “true” trend, which allows its sign to be correctly estimated by the OLS method. From Eq. (6), this is equal to mean that once the magnitude of the “estimated” trend is greater than the theoretical threshold, its sign can be guaranteed as the true sign of the “true” linear trend.

2.4 Comparing the theoretical threshold and the other criteria

In climate research, the de-trended time series of a given variable, such as SST, is generally dominated by a few modes. For example, the de-trended SST in the eastern equatorial Pacific is dominated by internal variation at the interannual, decadal and interdecadal time scales (as will be shown in Fig. 3). Therefore, it is reasonable to assume the de-trended time series takes the form of Eq. (3). In this case, the uncertainty in the sign of the “estimated” linear trend determined by the theoretical threshold will be more reliable than those determined by the confidence level and the criterion shown in W98 which assume the de-trended data as the white noise on the AR1 process. In addition, unlike the criterion proposed in L08, the theoretical threshold can be easily estimated in practice.

To show the advantage of the theoretical threshold, we analyze the “estimated” linear trend of the following equation:

$$y(t) = 0.1t + 6.58 \sin(0.2t + \varphi). \quad (12)$$

The parameters are chosen so that the last point with the “estimated” trend will be equal to the critical length (Eq. 9) when the phase is set to $\phi = 0$ (Eq. 16 in the appendix). Note that using other parameters will give a comparable conclusion.

Figure 2 gives the “estimated” trend of Eq. (12) as a function of the record length. As suggested by Eq. (5), the “estimated” trend will nonlinearly approach the “true” trend with increasing record length. Only when the record length is more than the critical length of 62.8, which is estimated by Eq. (9), the sign of the “estimated” trend will be the same as that of the “true” linear trend, which is positive. In practice, the “true” linear trend is unknown. Therefore, if using the criterion of $\|r(L)\| > B_t$ to justify the sign of the “estimated” trend, the record length needed is 72.

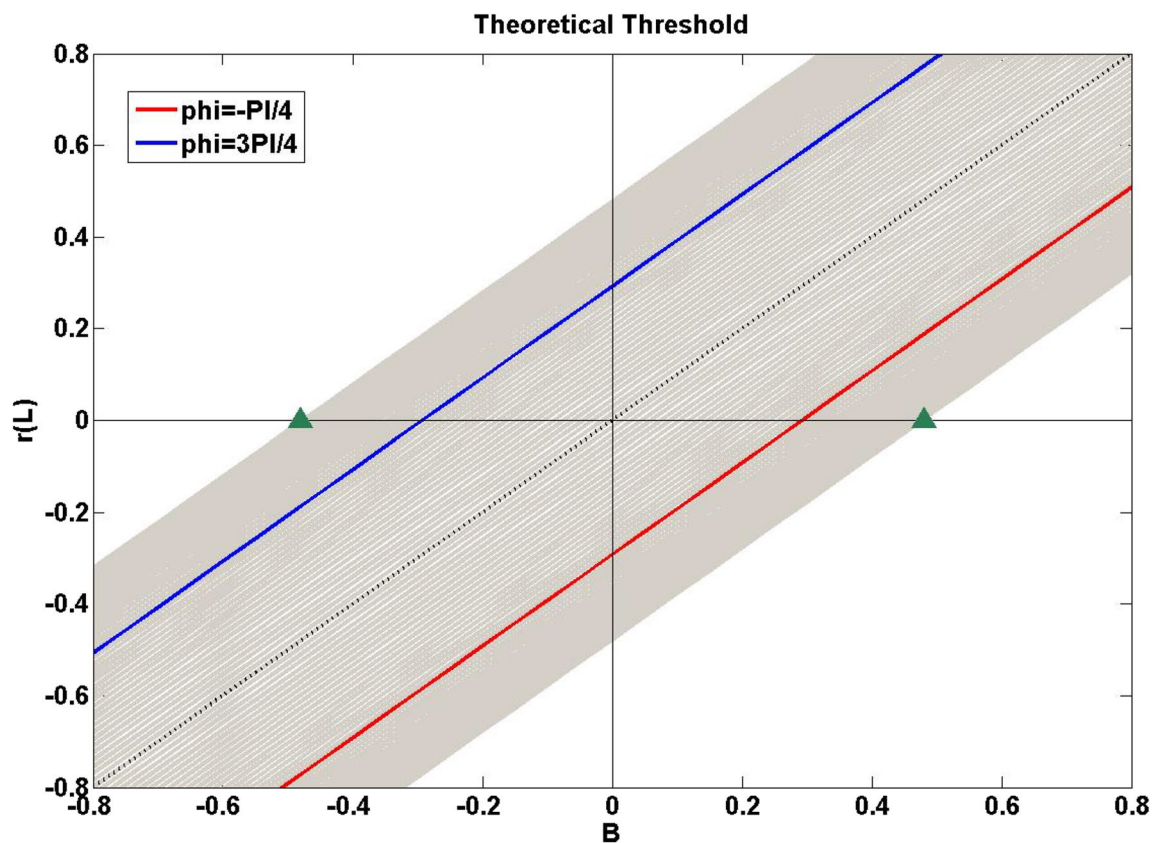


Fig. 1 The “estimated” linear trend $r(L)$ as a function of the “true” linear trend B . Red and blue lines denote the “estimated” linear trends when the phase is set as $-\frac{\pi}{4}$ and $\frac{3\pi}{4}$, respectively. Grey shading

denotes the span of $r(L)$ when using all possible phases from $-\pi$ to π . The black dashed line denotes $r(L) = B$. The green triangles denote the theoretical thresholds

However, using other criteria will get an unreliable estimation of the “true” linear trend. Suppose that one only has the record with a length of 30, which is shorter than the required length estimated by the new criterion. The “estimated” linear trend for this short record is -0.350 , and the 95% confidence interval is $[-0.362, -0.339]$. If one justifies the “estimated” linear trend by the 95% confidence test, it will be concluded that the underlying trend in Eq. (12) is a significant negative value. The value of U_{W08} and U_{L08} are 0.054 and 0.048, respectively, both smaller than the magnitude of the “estimated” trend. As a result, using the criteria proposed in W98 and L08 will lead one to believe that the linear trend underlying in Eq. (12) is negative, which is opposite to the sign of the “true” linear trend (0.1). On the other hand, the required theoretical threshold proposed here is 0.439 when the record length is 30, larger than the magnitude of the “estimated” trend. Therefore, the resulting negative value (-0.350) cannot be fully interpreted as the underlying trend in Eq. (12). This example clearly indicates that the criterion proposed here is more reliable for justifying the sign of the “estimated” trend when the time series is dominated by the oscillatory term. It should be noted that

since the autocorrelation at lag 1 is very close to 1.0 for Eq. (12), U_{W08} is estimated by the more complex formula shown in the appendix in W98, instead of Eq. (1). To estimate U_{L08} , following Phojanamongkolkij et al. (2014), the autocorrelation coefficient is set up to $10\log_{10}L$ lags.

3 Effect of internal variation on long-term linear trend in global SST

As an application of the criterion proposed in Sect. 2, we now investigate the uncertainty of long-term linear trend in global SST in the recent history. The reconstruction is HadISST1 (Rayner et al. 2003), one of the widely used datasets for climate research. The period is from 1881 to 2013, a total of 1596 monthly records. The SST anomaly is obtained by removing the climatological seasonal mean.

In order to estimate the theoretical threshold for the global SST, we need to know the amplitudes and frequencies for the internal variation at each grid. A widely used method is the fast Fourier transform (FFT). As an example, Fig. 3a gives the time series of SST anomaly in the

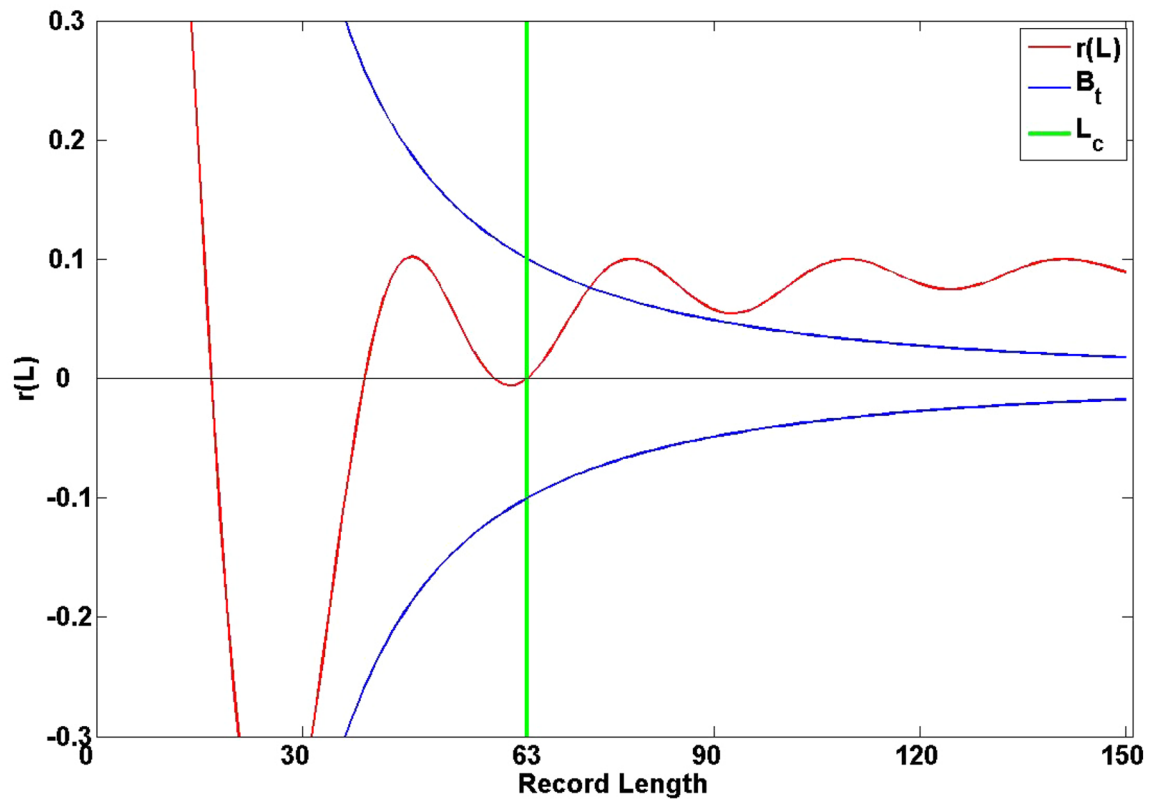


Fig. 2 The “estimated” linear trend (red curve) and the range of the theoretical threshold (blue curve) as a function of record length. The vertical green line denotes the critical length

equatorial Pacific at (120°W, 0°). The amplitudes and periods of all internal variation at this grid identified by the FFT are shown in Fig. 3b. The SST anomaly at this grid has two distinguished spectral peaks: one with periods from 2 to 10 years associated with the ENSO, and the other with a peak at 60 years related to the Pacific Decadal Oscillation (PDO; Mantua et al. 1997) and the Interdecadal Pacific Oscillation (IPO; Power et al. 1999). The spectral peak with periods from 10 to 20 years can also be explained by the PDO (Minobe 1997).

Because the record length of the monthly SST data used here is 1596, the longest 2-factorial period identified by the FFT algorithm is 2048 months. Since we cannot fully estimate the oscillation with a period of 2048 months by the 1596-month-long record, only the sinusoids with periods no longer than 1024 months are regarded as internal variation and are used in calculating the theoretical threshold according to Eq. (7). In this case, the frequency of the internal variation satisfies $\omega \geq \frac{2\pi}{L}$, where L equals 1596. The condition of $\omega L \gg 2$ required by Eq. (14) thus can be satisfied. The noise effect that denotes the uncertainty of “estimated” trend due to the white noise is estimated by Eq. (10). The variance of the noise term used in Eq. (10) is assumed as the variance of the time series for simplicity.

Figure 4a gives the theoretical threshold for the global SST trend. We can see that the maximum theoretical threshold with the magnitude of 1.75 °C per century is confined to the western boundary regions. Therefore, only when the magnitude of the “estimated” trend in these regions is larger than 1.75 °C per century, its sign is reliable. In the eastern tropical Pacific and the North Pacific, the required theoretical threshold is about 1.0 °C per century. In general, areas with large theoretical threshold are regions with strong decadal (10–30 years) variability, as implicated by Eq. (7). It is noted that the theoretical thresholds required by the North Pacific is comparable with that required by the eastern equatorial Pacific, the region where ENSO dominates. The reason is that there is strong decadal variability in the North Pacific (Fig. 4a), namely, the PDO (Mantua et al. 1997). The theoretical threshold depends on both amplitude and frequency of local internal variation (Eq. 7). While the amplitude of the decadal SST variability in the North Pacific is smaller than that of ENSO in the eastern equatorial Pacific, its lower frequency could let the North Pacific requires a theoretical threshold as large as that for the eastern tropical Pacific.

The ratio of the noise effect and the theoretical threshold is shown in Fig. 4b. It is clear that in most parts of

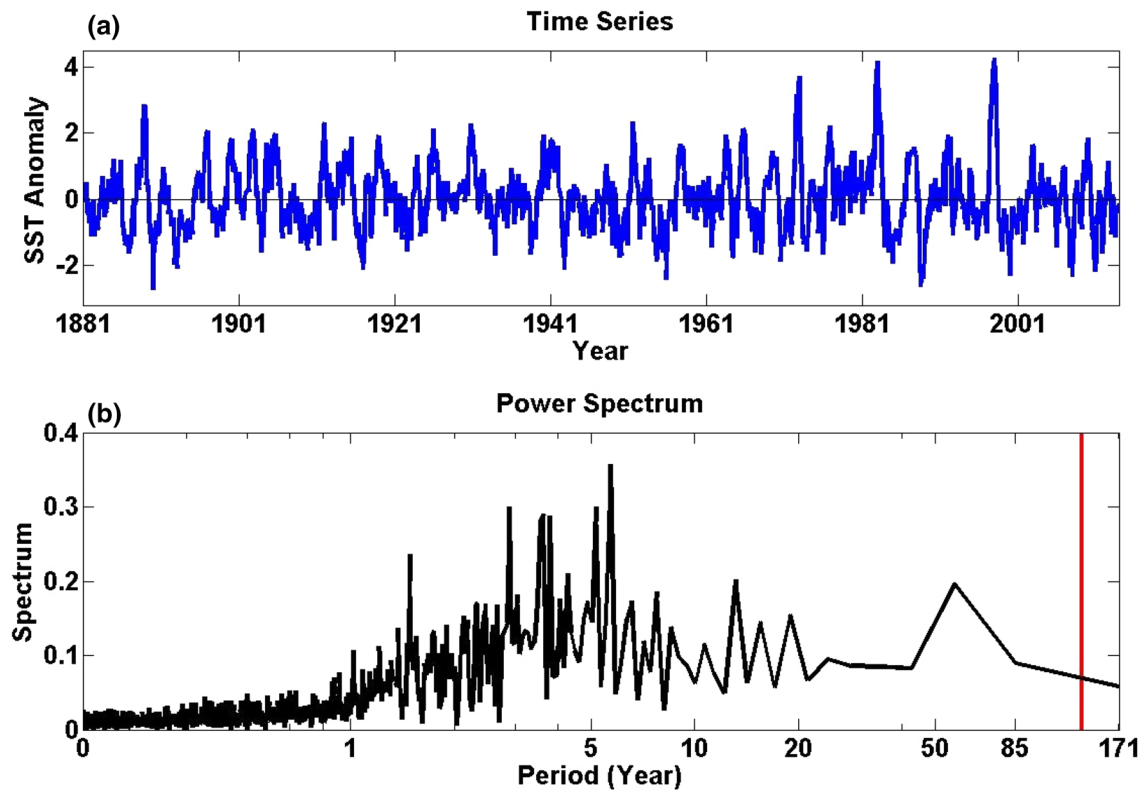


Fig. 3 **a** SST time series at (120°W, 0°). **b** The power spectrum of SST time series shown in **a** estimated by the FFT analysis. The vertical red line in **b** denotes the record length of 133 years

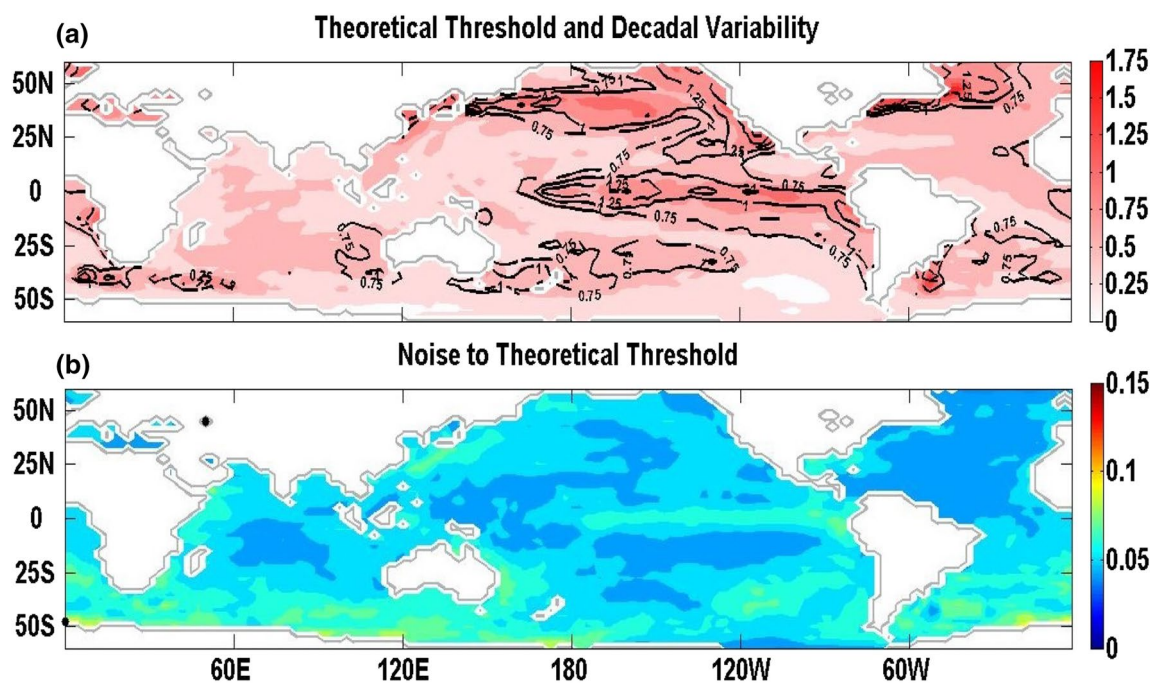


Fig. 4 **a** The theoretical threshold (color, unit of °C per century) and the decadal (10–30 years) variability (contour, unit of °C) of global SST linear trend. **b** The ratio of the noise effect and the theoretical

threshold. In **a**, only areas with SST decadal variability larger than 0.5 °C are shown for neat. The black dots in **b** denote where the ratio is larger than 0.1

the global ocean, except for the Mediterranean region, the noise effect counts <10% of the uncertainty of “estimated” trend due to internal variation. We thus could neglect the uncertainty due to the noise term at a good approximation. In this case, the uncertainty of “estimated” linear trend in global SST can be measured by the theoretical threshold.

Figure 5 presents the linear trend in global SST from 1881 to 2013. In most parts of the global ocean, the SST exhibits a warming trend. In the eastern equatorial Pacific and North Atlantic, the trend is negative (cooling). However, the magnitudes of SST trend in most parts of these regions (dotted regions) are smaller than the required theoretical threshold shown in Fig. 4a. Therefore, the positive or/and negative “estimated” trends in these regions may depend on the phase of local multi-scale variation, and thus cannot be fully interpreted as the result of the anthropogenic forcing. Only in the western boundary regions, the tropical Indian Ocean, part of the South Atlantic, and the region between the southern Indian Ocean and the Antarctica, the warming trends found can be treated as the underlying long-term change. Note that except the western boundary regions, most of areas with significant warming trend are regions with relatively weak decadal SST variability (Fig. 4a). The strong warming trend over the western boundary regions is argued to be associated with a synchronous poleward shift and/or intensification of global subtropical western boundary currents in conjunction with a systematic change in winds over both hemispheres (Wu

et al. 2012). Du and Xie (2008) suggested that the increased surface relative humidity and stability acted to reduce local evaporation and further amplified SST warming in the twentieth century. Details on the formation of the global warming SST pattern can be found in Xie et al. (2010).

Assuming the “true” linear trends are those “estimated” trends shown in Fig. 5, we can now estimate the critical length required by global SST. By the definition of the critical length (Eq. 9), it is equivalent to estimate how long the record needs to be to ensure that the sign of the known “true” linear trend that would not be incorrectly estimated by the OLS method. Figure 6 shows the critical length required by global SST. Near the western boundary regions, since the magnitude of the trend is large, records of several decades will be adequate to detect its sign. Also in the South Atlantic, tropical Indian Ocean and the Southern Ocean, the required critical lengths are within the length of the available reconstructions (133 years here). However, in the central southern Indian Ocean, the eastern off-equatorial Pacific and a large area of the North Pacific, one needs a record of about 300 years to ensure the sign of the trend. Apparently, this record length is much longer than the length of currently available reconstructions. For the regions such as the eastern equatorial Pacific, the South Pacific near the Antarctica, the central North Pacific and the North Atlantic, the record length needed to exclude the uncertainty due to multi-scale internal variation is more than 1000 years. For example, the “estimated” SST trend

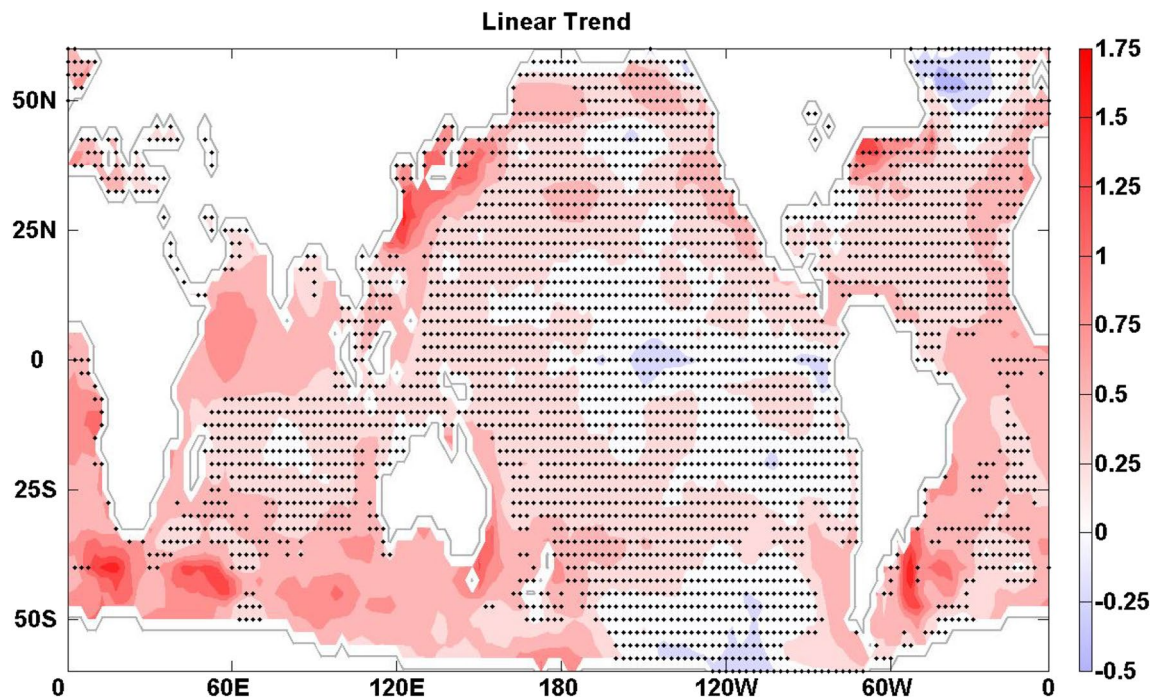


Fig. 5 The linear trend in global SST (°C per century) from 1881 to 2013. Each *black dot* denotes where the magnitude of the linear trend is smaller than the theoretical threshold

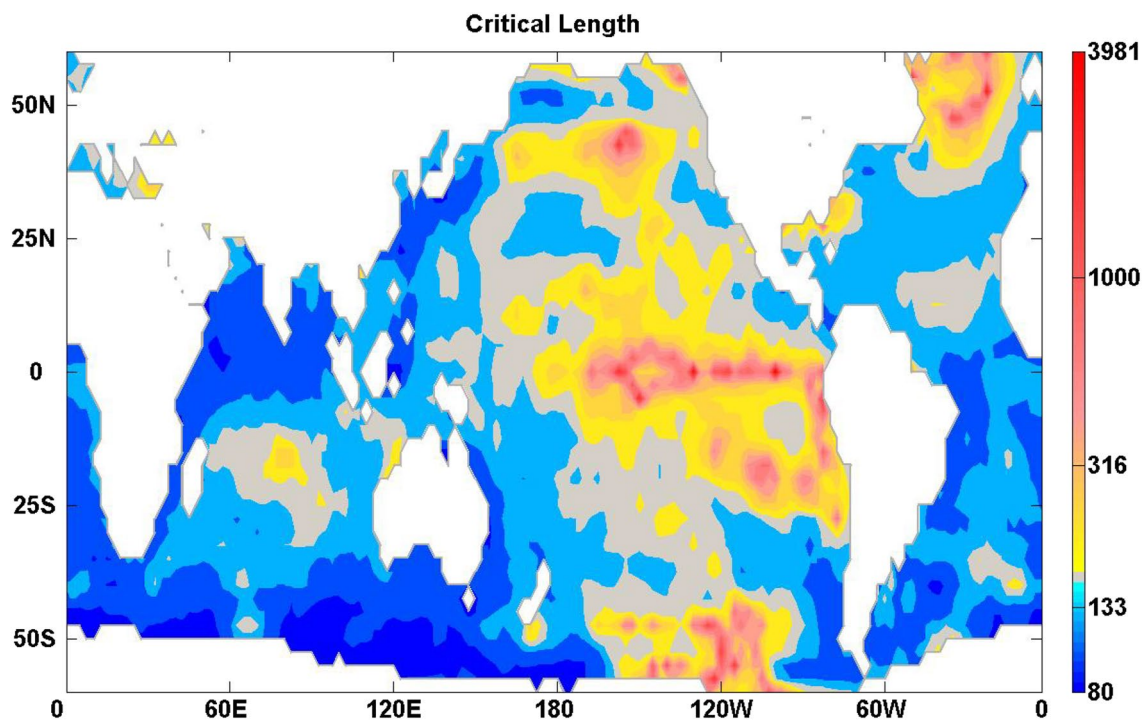


Fig. 6 The critical length (units: year) of global SST linear trend

near (120°W, 60°S) is about -0.5°C per century (Fig. 5). If we are to use this “estimated” trend to represent the “true” trend there, our estimation should be based on a 2000-year record (Fig. 6).

Above analyses indicate that the sign of the “estimated” trend in most parts of the global SST is sensitive to the time period used, and thus it may not be a robust trend. We now discuss whether global warming in the recent history is a robust positive trend or not. Figure 7a presents the evolution of global-mean SST anomaly. It is clear that the global-mean SST anomaly exhibits strong interannual variation. Also seen is that there were segments with a sharp warming (for example, periods of 1911–1941 and of 1975–2000) and with a weak warming or slowed-down warming (for example, 1900–1910 and 1940–1950), indicating the internal variation of global-mean SST anomaly at the decadal and interdecadal time scales.

Figure 7b gives the “estimated” trend of global-mean SST anomaly as a function of the starting year. The ending year used to calculate the “estimated” trend is fixed at 2013. The corresponding theoretical threshold and noise effect, which depend on the record length, are also shown. The “estimated” trend in global-mean SST anomaly is negative or close to zero when the starting year is after 1998, known as the global warming hiatus (Trenberth and Fasullo 2013). With increasing record length, the “estimated” trend in global-mean SST anomaly gradually increases and reaches the maximum of 0.93°C per century when the starting year

is around 1974, the beginning of the second sharp increasing segment in global-mean SST anomaly (Fig. 7a). When the starting year is shifted to an earlier year, the “estimated” linear trend becomes stationary with a magnitude of about 0.5°C per century.

To clarify whether the sign of the “estimated” trend is robust, we need to compare the magnitude of the “estimated” trend with the required theoretical threshold. Note that the noise effect is very small for global-mean SST anomaly. Because of the strong internal variation in global-mean SST anomaly, especially the variation at the interdecadal time scale, the theoretical threshold is very large when the record length is short. For example, when the record length is 40 years, the theoretical threshold is 2.55°C per century. Because of this, we cannot use the finding that the 40-year (from 1974 to 2013) trend is 0.93°C per century to argue that the long-term trend in global-mean SST anomaly was warming, just as we cannot use the negative trend of -0.11°C per century based on another 40-year record (from 1937 to 1976) to argue that the long-term trend in global-mean SST anomaly was cooling.

However, the theoretical threshold sharply decreases with the square of record length. When the starting year is set before 1924, the magnitude of the positive “estimated” trend will be larger than the corresponding theoretical threshold. Therefore, the warming trend in global-mean SST anomaly can be guaranteed only if the record length is longer than 90 years (1924–2013). Otherwise, namely,

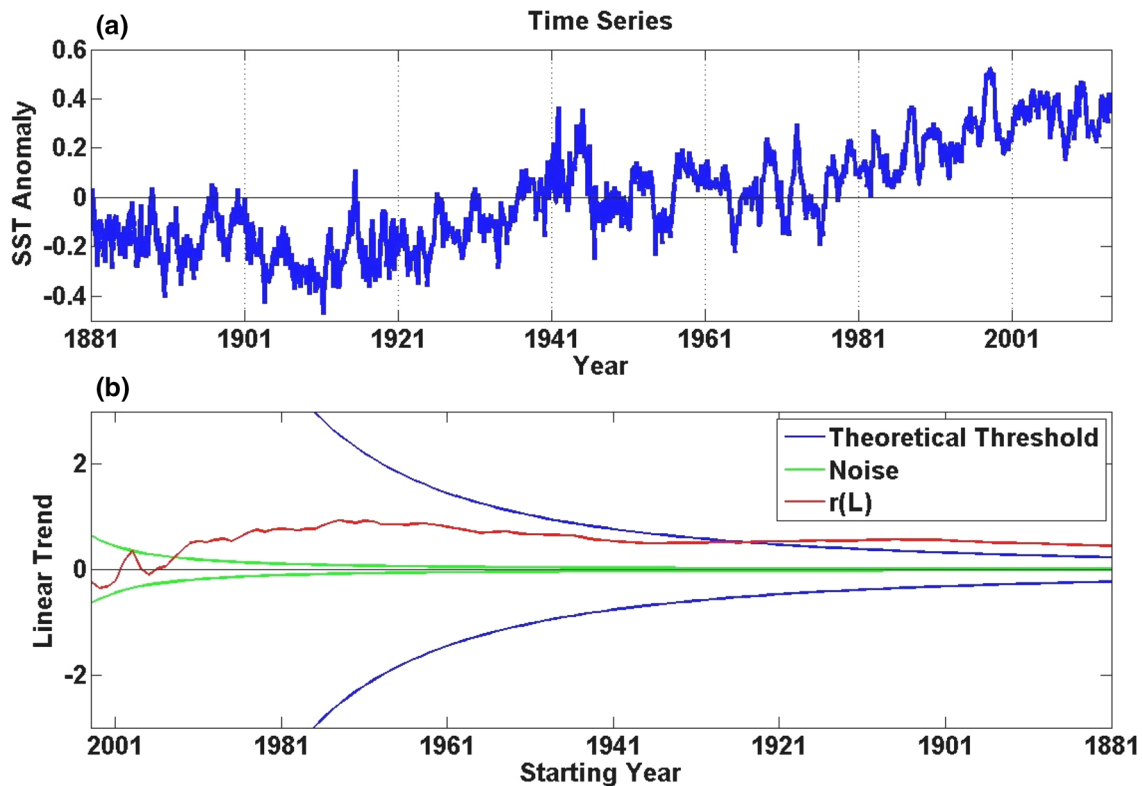


Fig. 7 **a** Global-mean SST anomaly. **b** The range of theoretical threshold (blue line), the noise effect (green) and the “estimated” linear trend (red line) as a function of the starting year. The ending year is fixed at 2013

when the record length is shorter than 90 years, the sign of the resulting estimated trend may depend on the time period used, just as the case study of the opposite trend based on two 40-years records given above.

As another application of the new criterion proposed here, we explore the contribution of multi-scale variation on the recent global warming hiatus, which started from the early 2000s. The linear trend of global-mean SST in the period of 2000–2013 was 0.148°C per century. On the other hand, the uncertainties associated with internal variation at interannual (roughly 2–10 years), decadal (10–30 years), multi-decadal (30–70 years), and centennial (more than 70 years) time scales are 0.092, 0.072, 0.176, and 0.069°C per century, respectively. Apparently, the combined uncertainty due to the interannual and decadal variations ($0.092 + 0.072 = 0.164$) is larger than the magnitude of the recent hiatus. Therefore, variations at the interannual and decadal time scales are large enough to explain the recent hiatus. The leading modes of global SST at the interannual and decadal time scales are ENSO and PDO, respectively. Many recent studies (e.g., Kosaka and Xie 2013) have shown that the cooling SST in the eastern equatorial Pacific, which is closely related to the negative phase of PDO, can lead to the hiatus.

4 Conclusion and discussion

In this study, we first carried out a theoretical analysis of the linear trend derived by the OLS method. It was found that the uncertainty of “estimated” linear trend consists of two terms: one related to multi-scale internal variation, and the other related to noise. A new theoretical threshold was then introduced to scale the uncertainty in the sign of the “estimated” trend due to the influence from multi-scale internal variation. When the amplitude and frequency of the multi-scale internal variation are given, the sign of the “estimated” trend can be determined for the underlying long-term change once its magnitude is greater than the theoretical threshold. Otherwise, the sign of the “estimated” trend may depend on the phase of the internal variation, namely, the underlying long-term trend may depend on the time period used. Comparing with other widely used criteria, we illustrated that, for the time series that is dominated by the oscillatory term, the criterion proposed in this study is more reliable for justifying the uncertainty in the sign of the “estimated” trend.

We then used this new criterion to discuss the uncertainty of the long-term linear trend in global SST from 1881 to 2013. It was found that the “estimated” warming trends in

most parts of the global ocean and the “estimated” cooling trends in the eastern equatorial Pacific and North Atlantic may not represent the signs of the actual underlying long-term changes related to the anthropogenic forcing since their magnitudes are smaller than the required theoretical thresholds. Only in the western boundary regions, the tropical Indian Ocean, part of the South Atlantic, and the region between the southern Indian Ocean and the Antarctica, the “estimated” warming trends found using the 133-year record can represent the true underlying long-term change. We also used this criterion to explore the uncertainty of global-mean SST anomaly in the recent history. We concluded that the global warming since 1881 is a robust trend. Moreover, by comparing the magnitude of global-mean SST trend since 2000 and the uncertainty in trend due to oscillations at different time scales, we suggested the variations at the inter-annual and decadal time scales are large enough to explain the recent global warming hiatus.

The methods proposed in W98 and L08 were designed to detect climate trends. The main difference between W98/L08 and this study resides in the form of the linear trend model. In W98, the de-trended part of a time series is assumed to follow the AR1 process. Although the uncertainty of the resulting trend can be easily estimated, any departure from the AR1 model will cause the method to perform poorly. While L08 method does not require any form of de-trended part of a time series, the difficulty in estimating autocorrelation coefficients at the infinite lag makes it hard to be used in practice. On the other hand, since distinct multi-scale variation does exist in climate variables, the linear trend model used in the current study is probably closer to the reality. The oscillatory information can be assessed by spectral analysis, making it easy to use in practice. In addition, the model used in this study allows us to estimate the contribution of various oscillatory terms to the uncertainty in trend. For example, we can estimate how strong ENSO and PDO will affect SST trend in the eastern equatorial Pacific, respectively. We will address these issues in our future study.

It should be highlighted here that applying this new criterion to the real data needs a good estimation of the oscillatory component in the time series. Although the FFT analysis is widely used, like any available spectral analysis, it has difficulty in fully assessing the oscillatory components with periods near the record length, or longer. As seen in Eq. (7), the theoretical threshold is a linear combination of oscillatory amplitudes. The confidence level of the theoretical threshold is therefore scaled by the bound of oscillatory amplitudes. As a result, a small deviation of the low-frequency oscillatory amplitude could lead to drastic impacts on the resulting trend (e.g., Harrison and Chiodi 2014). While the current study focused on the superiority of the new criterion in scaling the uncertainty in trend over the existing methods, cautions should be placed on

applying this criterion to the real data, especially on the estimation of low-frequency oscillatory component in the data.

Also noted here is the assumption of the white noise in the trend model used here. While we assume the autocorrelation in the data can be completely accounted for by the oscillatory term in the model, observed variability not accounted for by the oscillatory term could be red noise or multiplicative noise rather than white noise. Wunchsh (1999) pointed out the serious problem with statistical tests that assume the white noise type behavior when autocorrelation in fact exists. A more sophisticated model should consider, for example, the AR1 process (W98) as the noise term. The uncertainty due to the noise can then be estimated by Eq. (1) in this case. However, it should be kept in mind that no matter what kind of noise is assumed, there will always be uncertainty with the real data.

In this paper, we argued that large uncertainty of the long-term linear trend in global SST exists based on the HadISST1 reconstruction, which is a widely used interpolated archive. However, as pointed out in many studies, different datasets will give controversial estimations of the sign of the linear trend in regional SST (e.g., Karnauskas et al. 2009; Deser et al. 2010; Tokinaga et al. 2012). In addition, it is suggested that the interpolated archives contain large uncertainty owing to the poor data quality and data assimilation schemes in the record before the early twentieth century (Deser et al. 2010). Applying our new criterion to various SST datasets, especially the un-interpolated archives, will undoubtedly provide a better evaluation of the uncertainty of long-term trend in global SST.

Finally, all our analyses are based on the linear regression model. In reality, the long-term global SST change may contain nonlinear components. Some studies suggested that anthropogenic forcing could lead to decadal SST fluctuations from combined effects of different forcing factors (e.g., Meehl et al. 2009; Ting et al. 2009). In such case, the nonlinear analytical tools will be needed (Wu and Huang 2009). However, when carrying out research within the linear framework, it is recommended that one compare the magnitude of the “estimated” linear trend with its theoretical threshold. If the magnitude of the “estimated” trend is smaller than the corresponding theoretical threshold, one should be cautious when interpreting the sign of the “estimated” trend as the sign of the underlying long-term trend in the record used.

Acknowledgements The author thanks all three anomalous reviewers for their valuable comments. This work is supported by grants from the National Natural Science Foundation of China (41506025, 41690121, 41690120, 41476021 and 41321004), the National Program on Global Change and Air-Sea Interaction (GASI-IPOVAI-04 and GASI-IPOVAI-06), the Project of State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography

(SOEDZZ1504), and the Zhejiang Provincial Natural Science Foundation of China (LQ15D060004). Dr. Dake Chen and Dr. Zheqi Shen gave helpful comments during the course of this work. The author wishes to thank Dr. Zuojun Yu for improving the representation.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

$$r(O) = \frac{L \int_0^L t [\sum A_i \cdot \sin(\omega_i t + \varphi_i)] dt - \left(\int_0^L t dt \right) \left(\int_0^L [\sum A_i \cdot \sin(\omega_i t + \varphi_i)] dt \right)}{L \int_0^L t^2 dt - \left(\int_0^L t dt \right)^2} \quad (14)$$

$$= \frac{6}{L^3} \sum \frac{A_i}{\omega_i^2} \{ 2[\sin(\omega_i L + \varphi_i) - \sin(\varphi_i)] - \omega_i L [\cos(\omega_i L + \varphi_i) + \cos(\varphi_i)] \}$$

Appendix 1

See Fig. 8.

Appendix 2

For a continuous time series of $y = y(t)$, where $t \in [0, L]$, the linear trend is:

$$r = \frac{L \int_0^L t y(t) dt - \left(\int_0^L t dt \right) \left(\int_0^L y(t) dt \right)}{L \int_0^L t^2 dt - \left(\int_0^L t dt \right)^2} \quad (13)$$

$y(t)$ by $\sum A_i \cdot \sin(\omega_i t + \varphi_i)$, the part of the “estimated” linear trend related to the oscillatory term is:

When $\omega_i L$ is much > 2 , the first two terms in the bracket will be small and can be neglected. Therefore, we have:

$$r(O) \cong -\frac{6}{L^2} \sum \frac{A_i}{\omega_i} [\cos(\omega_i L + \varphi_i) + \cos(\varphi_i)] \quad (15)$$

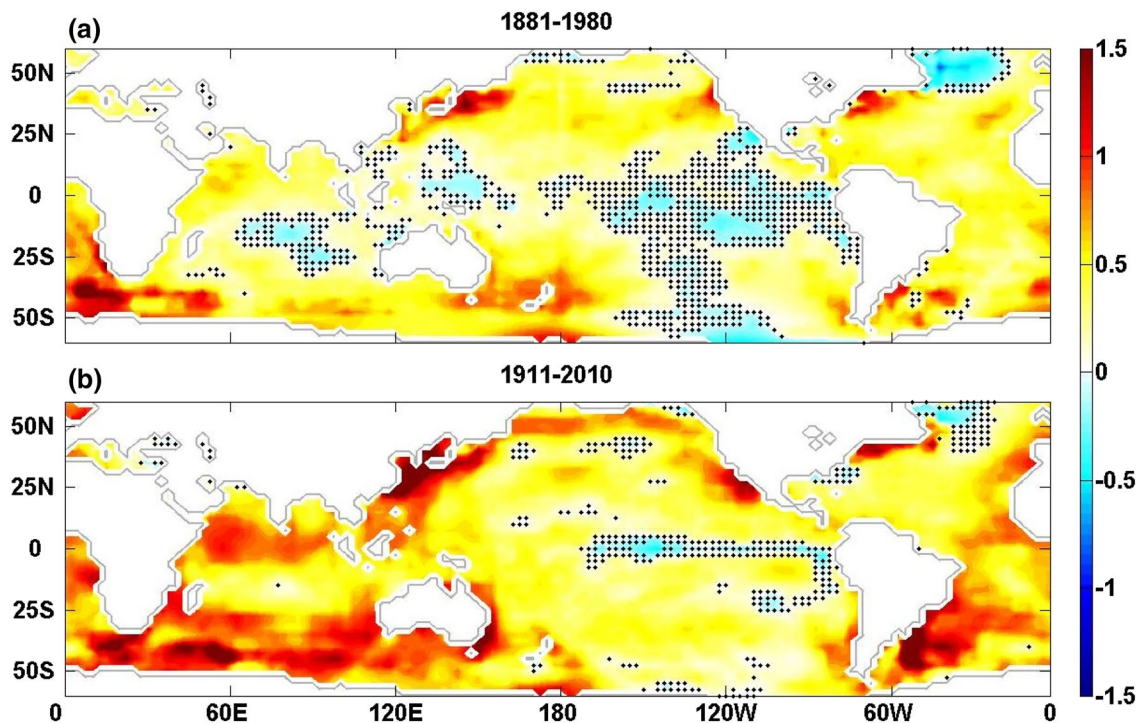


Fig. 8 The linear trend in global SST (°C per century) estimated by the ordinary least-square (OLS) method over 100 years from 1881 to 1980 (a) and from 1911 to 2010 (b). Each black dot denotes where the linear trend coefficient is insignificant from zero at the 95% confidence level

To make the last point with the “estimated” trend equal the critical length (Eq. 9), the sum in the bracket in Eq. (16) should be equal to -2 or 2 . Thus, we need:

$$\begin{cases} \varphi_i = k_1 \pi \\ \omega_i L_C = k_2 \pi \\ L_C = \sqrt{\frac{12}{B} \sum \frac{A_i}{\omega_i}} \end{cases} \quad (16)$$

In the idealized case shown in Sect. 2.4 (Eq. 12), k_1 and k_2 are set as 0 and 4, respectively.

References

- Cai W et al (2013) Increasing frequency of extreme El Niño events due to greenhouse warming. *Nat Clim Change* 4:111–116
- Deser C, Phillips AS, Alexander MA (2010) Twentieth century tropical sea surface temperature trends revisited. *Geophys Res Lett* 37:L10701. doi:[10.1029/2010GL043321](https://doi.org/10.1029/2010GL043321)
- Du Y, Xie S-P (2008) Role of atmospheric adjustments in the tropical Indian Ocean warming during the 20th century in climate models. *Geophys Res Lett* 35:L08712
- Harrison DE, Chiodi AM (2014) Multi-decadal variability and trends in the El Niño–Southern Oscillation and Tropical Pacific fisheries implications. *Deep-Sea Res II* 113:9–21
- Karnauskas KB, Seager R, Kaplan A, Kushnir Y, Cane MA (2009) Observed strengthening of the zonal sea surface temperature gradient across the equatorial Pacific Ocean. *J Clim* 22:4316–4321
- Kosaka Y, Xie S-P (2013) Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature* 501:403–407
- L’Heureux ML, Lee S, Lyon B (2013) Recent multidecadal strengthening of the Walker circulation across the tropical Pacific. *Nat Clim Change* 3:571–576
- Leroy SS, Anderson JG, Ohring G (2008) Climate signal detection times and constraints on climate benchmark accuracy requirements. *J Clim* 21:841–846
- Loeb NG et al (2007) Multi-instrument comparison of top-of-atmosphere reflected solar radiation. *J Clim* 20:575–591
- Luo J-J, Sasaki W, Masumoto Y (2012) Indian Ocean warming modulates Pacific climate change. *Proc Natl Acad Sci USA* 109(46):18701–18706
- Mantua NJ, Hare SR, Zhang Y, Wallace JM, Francis RC (1997) A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull Am Meteorol Soc* 78:1069–1079
- Meehl GA, Hu A, Santer BD (2009) The mid-1970s climate shift in the Pacific and the relative roles of forced versus inherent decadal variability. *J Clim* 22:780–792
- Minobe S (1997) A 50–70 year climatic oscillation over the North Pacific and North America. *Geophys Res Lett* 24:683–686
- Phojanamongkolkij N, Kato S, Wielicki BA, Taylor PC, Mlynarczyk MG (2014) A comparison of climate signal trend detection uncertainty methods. *J Clim* 27:3363–3376
- Power S, Casey T, Folland C, Colman A, Mehta V (1999) Interdecadal modulation of the impact of ENSO on Australia. *Clim Dyn* 15:319–324
- Rayner NA et al (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J Geophys Res* 108(D14):4407. doi:[10.1029/2002JD002670](https://doi.org/10.1029/2002JD002670)
- Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experimental design. *Bull Am Meteorol Soc* 93:485–498
- Ting M, Kushnir Y, Seager R, Li C (2009) Forced and internal twentieth-century SST trends in the North Atlantic. *J Clim* 22(6):1469–1481. DOI:[10.1175/2008JCLI2561.1](https://doi.org/10.1175/2008JCLI2561.1)
- Tokina H, Xie S-P, Deser C, Kosaka Y, Okumura TM (2012) Slowdown of the Walker circulation driven by tropical Indo-Pacific warming. *Nature* 491:439–443
- Trenberth KE, Fasullo JT (2013) An apparent hiatus in global warming? *Earth’s Future* 1:19–32. doi:[10.1002/2013EF000165](https://doi.org/10.1002/2013EF000165)
- von Storch H, Zwiers F (1999) Statistical analysis in climate research. Cambridge University Press, Cambridge
- Weatherhead EC, Reinsel GC, Tiao GC, Meng X, Choi D, Cheang W, Keller T, DeLuisi J, Wuebbles DJ, Kerr JB, Miller AJ, Oltmans SJ, Frederick JE (1998) Factors affecting the detection of trends: Statistical considerations and applications to environmental data. *J Geophys Res* 103:17149–17161
- Wielicki BA et al (2013) Achieving climate change absolute accuracy in orbit. *Bull Am Meteorol Soc* 94:1519–1539. doi:[10.1175/BAMS-D-12-00149.1](https://doi.org/10.1175/BAMS-D-12-00149.1)
- Wu Z, Huang N (2009) Ensemble empirical mode decomposition: a noise assisted data analysis method. *Adv Adapt Data Anal* 1(01):1–41
- Wu L, Cai W, Zhang L, Nakamura H, Timmermann A, Joyce T, McPhaden MJ, Alexander M, Qiu B, Visbeck M, Chang P, Benjamin G (2012) Enhanced warming over the global subtropical western boundary currents. *Nat Clim Change* 2(3):161–166
- Wunsch C (1999) The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations. *Bull Am Meteorol Soc* 80:245–255
- Xie S-P, Deser C, Vecchi GA, Ma J, Teng H, Wittenberg AT (2010) Global warming pattern formation: sea surface temperature and rainfall. *J Clim* 23:966–986